National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Apache SDAP – A Disruptive Technology Solution for Earth Science

Thomas Huang
thomas.huang@jpl.nasa.gov

NASA AIST: OceanWorks/PI

Group Supervisor - Computer Science for Data-Intensive Applications
Strategic Lead - Interactive Data Analytics

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109-8099, U.S.A.

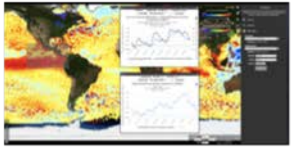
[CL # 19-3114]

Acknowledgement

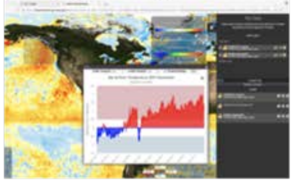
Ed Armstrong/JPL	Maya DeBellis/JPL	Joe Jacob/JPL	David Moroni/JPL	Rob Toaz/JPL
Jason Barnett/LARC	Rich Doyle/JPL	Zaihua Ji/NCAR	Kevin Murphy/NASA	Vardis Tsontos/JPL
Andrew Bingham/JPL	Jocelyn Elya/FSU	Yongyao Jiang/GMU	Charles Norton/JPL	Suresh Vannan/JPL
Carmen Boening/JPL	Ian Fenty/JPL	Felix Landerer/JPL	Jean-Francois Piolle/IFREMER	Jorge Vazquez/JPL
Mark Bourassa/FSU	Kevin Gill/JPL	Yun Li/GMU	Nga Quach/JPL	Ou Wang/JPL
Mike Chin/JPL	Frank Greguska/JPL	Eric Lindstrom/NASA	Brandi Quam/NASA	Brian Wilson/JPL
Marge Cole/NASA	Patrick Heimbach/UT Austin	Mike Little/NASA	Shawn Smith/FSU	Steve Worley/NCAR
Tom Cram/NCAR	Ben Holt/JPL	Chris Lynnes/NASA	Ben Smith/JPL	Elizabeth Yam/JPL
Dan Crichton/JPL	Thomas Huang/JPL	Lewis McGibbney/JPL	Adam Stallard/FSU	Phil Yang/GMU

Enabling Next Generation of Ocean Science Tools and Services

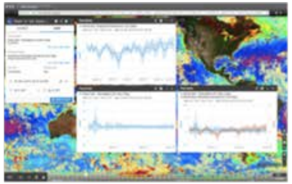
NASA Sea Level Change Portal



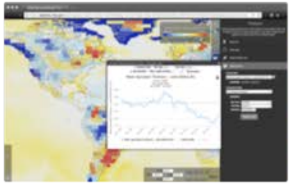
Oceanographic Anomaly Detection



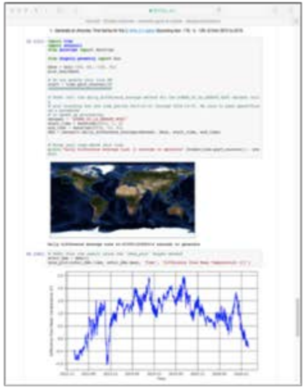
PO.DAAC State Of The Ocean



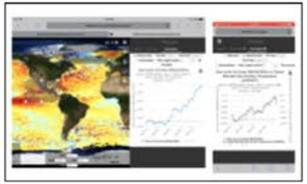
Hydrological Basin Analysis



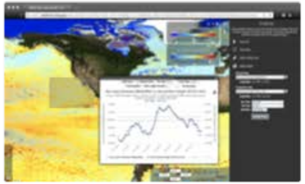
Jupyter Notebook - Interactive Workbench



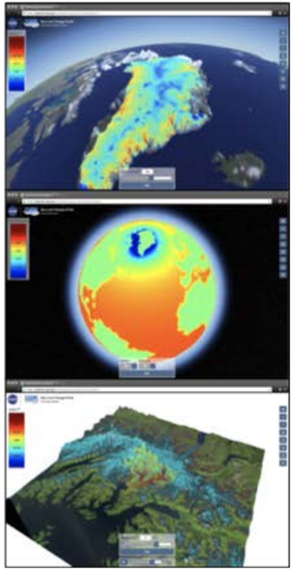
Mobile Analysis



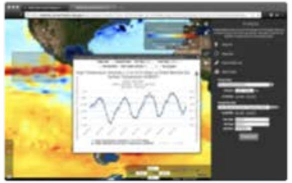
In Situ Data Analysis



Model Simulations



Model - Observation Comparison



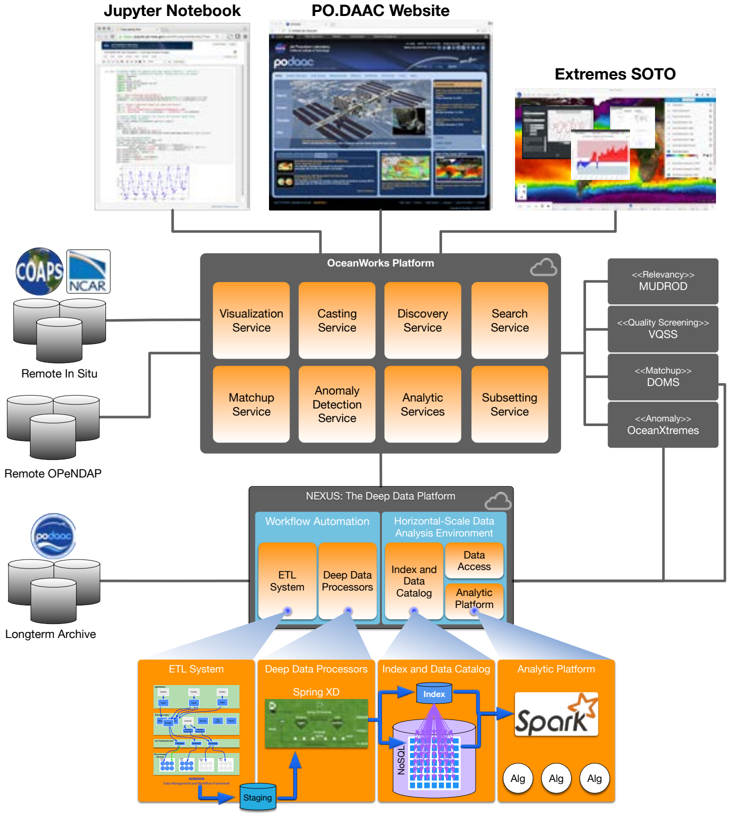
Integrated Search and Discovery

National Aeronautics and Space Administration

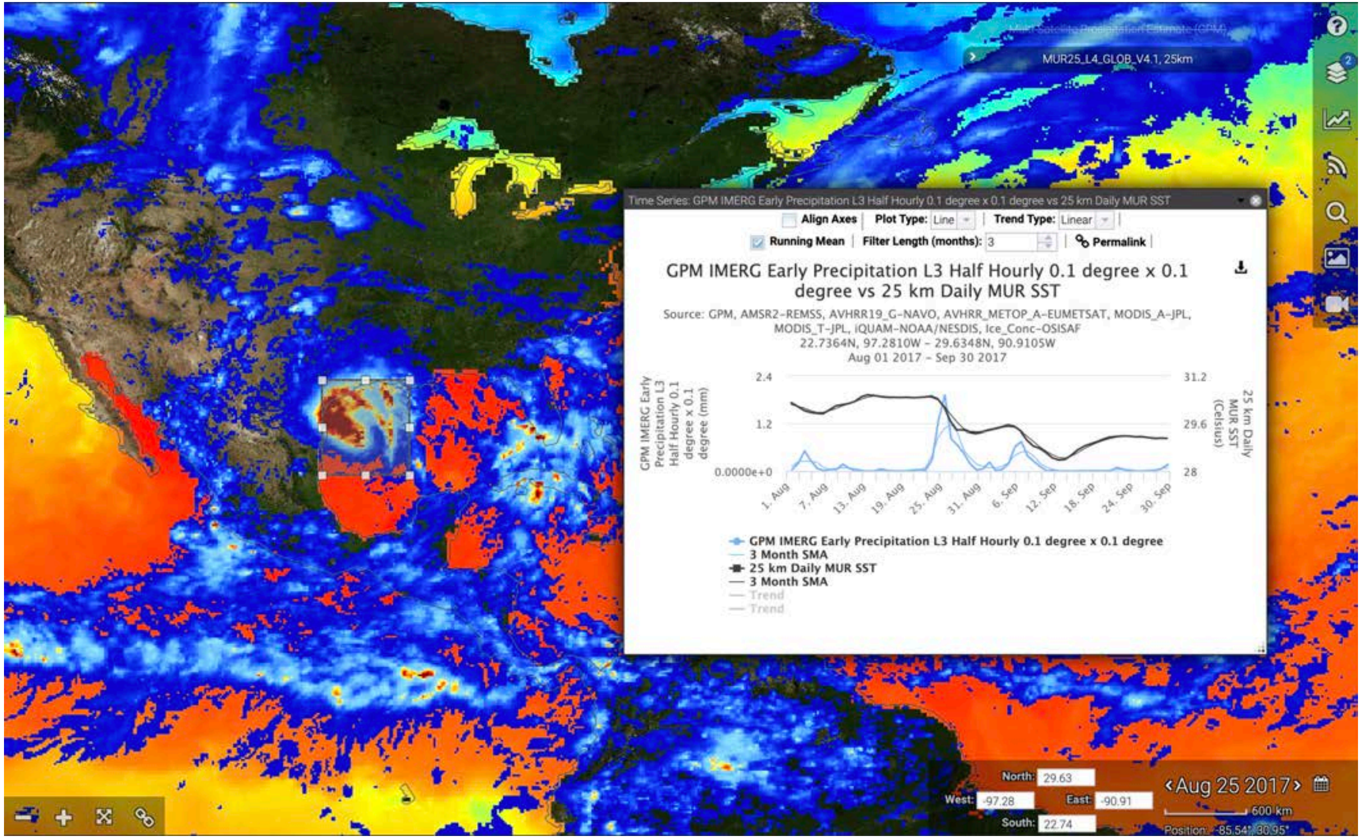
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

AIST OceanWorks: Ocean Science Platform

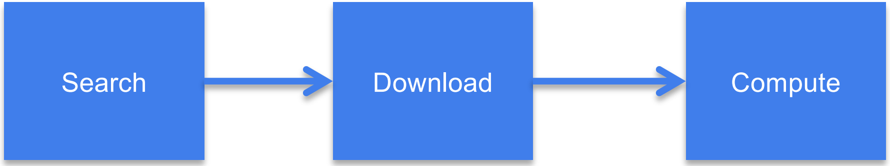
- **OceanWorks** is to establish an **Integrated Data Analytic Center Framework** at the NASA Physical Oceanography Distributed Active Archive Center (PO.DAAC) for Big Ocean Science
- Focuses on technology integration, advancement and maturity
- Collaboration between JPL, FSU, NCAR, and GMU
- Bringing together PO.DAAC-related big data technologies
  - Anomaly detection and ocean science
  - GIS-based big data analytic platform
  - Distributed in situ to satellite matchup
  - Search relevancy and discovery – linking datasets, services, and anomalies through recommendations
  - Metadata translation and services aggregation
  - Fast data subsetting
  - Virtualized Quality Screening Service

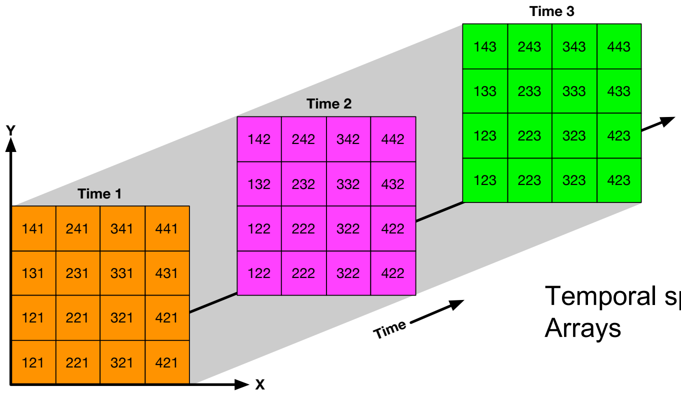


Analyze Hurricane Harvey using GPM and SST
Aug 17, 2017 – Sept. 2, 2017



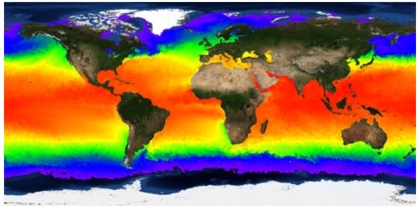
Traditional Method for Analyze Satellite Measurements

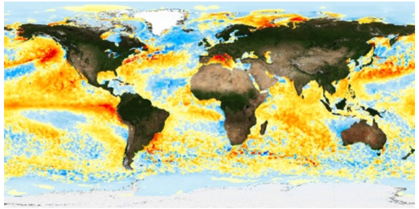


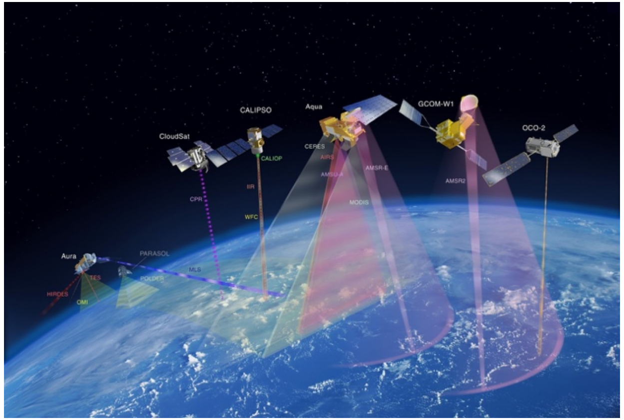


Temporal spatial Arrays

- Depending on the data volume (size and number of files)
- It could take many hours of download – (e.g. 10yr of observational data could yield thousands of files)
- It could take many hours of computation
- It requires expensive local computing resource (CPU + RAM + Storage)
- After result is produced, purge downloaded files







Observation

- Traditional methods for data analysis (time-series, distribution, climatology generation) can't scale to handle large volume, high-resolution data. They perform poorly
- Performance suffers when involve large files and/or large collection of files
- A high-performance data analysis solution must be free from file I/O bottleneck

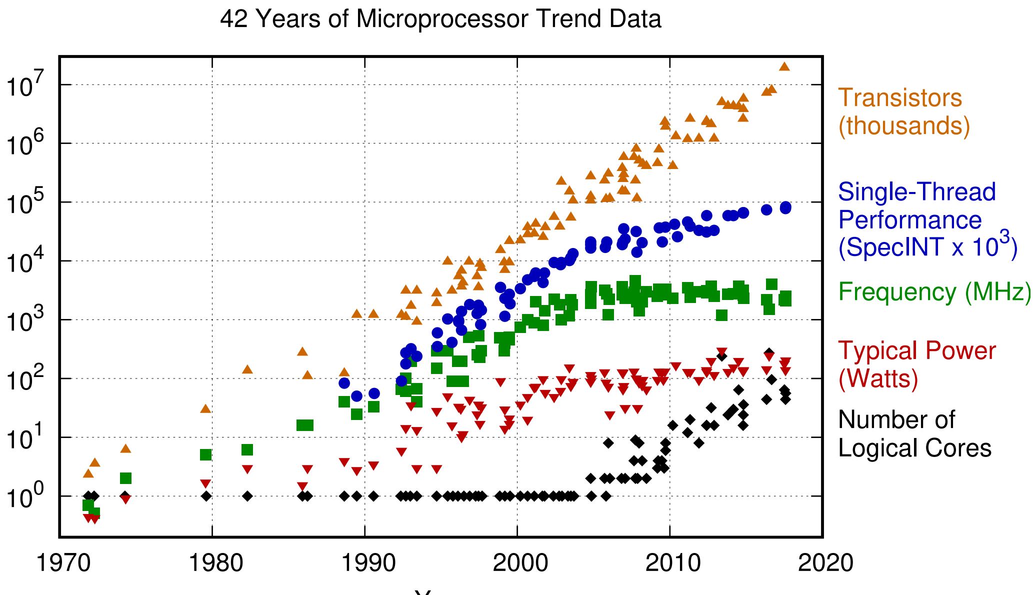
Processors are not Getting Faster

2004: First Pentium 4 processor with 3.0GHz clock speed

2018: Apple's MacBook Pro has clock speed of 2.7GHz

14 years later, not much has gain in raw processing power

Modern big data architects are required to "think outside of the box". Literally!



Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2017 by K. Rupp

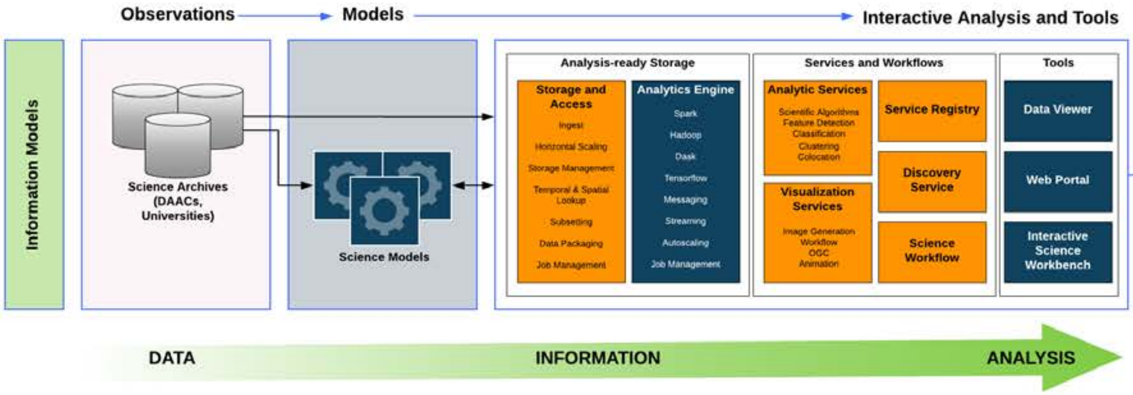
Disruptive Innovation

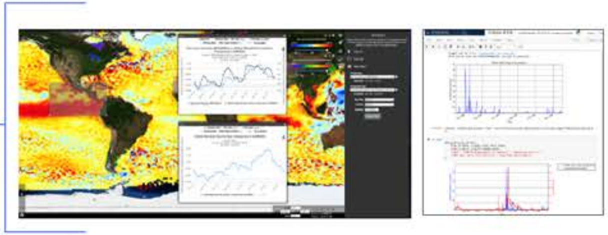
A process by which a product or service takes root initially in simply applications at the bottom of a market and then relentlessly moves up market, eventually displacing established competitors. - Clayton Christensen

products that require us to change our current mode of behavior or to modify our products and services we rely on. - Geoffrey Moore

Integrated Science Data Analytics Platform

Creating SaaS and PaaS for Science Tools and Services





Systematic analysis of massive data

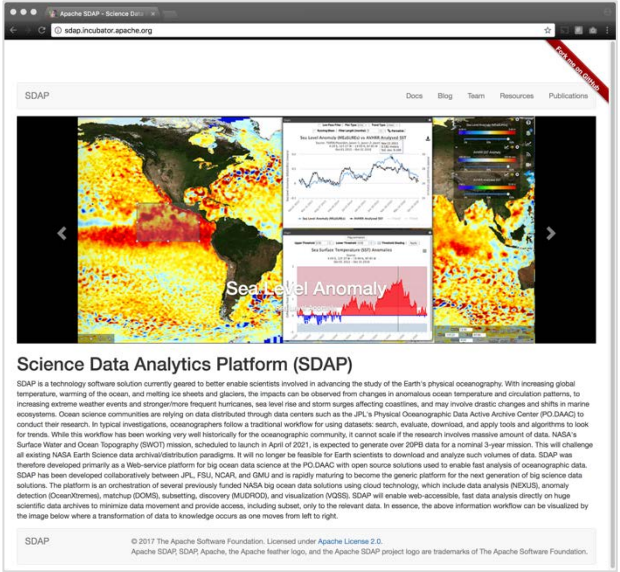
- **Integrated Science Data Analytics Platform**: an analytic center framework to provide an environment for conducting a science investigation
  - Enables the confluence of resources for that investigation
  - Tailored to the individual study area (physical ocean, sea level, etc.)
- Harmonizes data, tools and computational resources to permit the research community to focus on the investigation
- Scale computational and data infrastructures
- Shift towards integrated data analytics
- Algorithms for identifying and extracting interesting features and patterns

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Free and Open Source Software (FOSS)

- After more than two years of active development, on October 2017 we established Apache Software Foundation and established the **Science Data Analytics Platform (SDAP)** in the **Apache Incubator**
- Technology sharing through Free and Open Source Software (FOSS)
- Why? Further technology evolution that is restricted by projects / missions
- It is more than GitHub
    - Quarterly reporting
    - Reports are open for community review by over 6000 committers
    - SDAP has a group of appointed international mentors
- **SDAP and many of its affiliated projects are now being developed in the open**
    - Support local cluster and cloud computing platform support
    - Fully containerized using Docker and Kubernetes
    - Infrastructure orchestration using Amazon CloudFormation
    - Satellite and model data analysis: time series, correlation map,
    - In situ data analysis and collocation with satellite measurements
    - Fast data subsetting
    - Upload and execute custom parallel analytic algorithms
    - Data services integration architecture
    - OpenSearch and dynamic metadata translation
    - Mining of user interaction and data to enable discovery and recommendations



http://sdap.apache.org

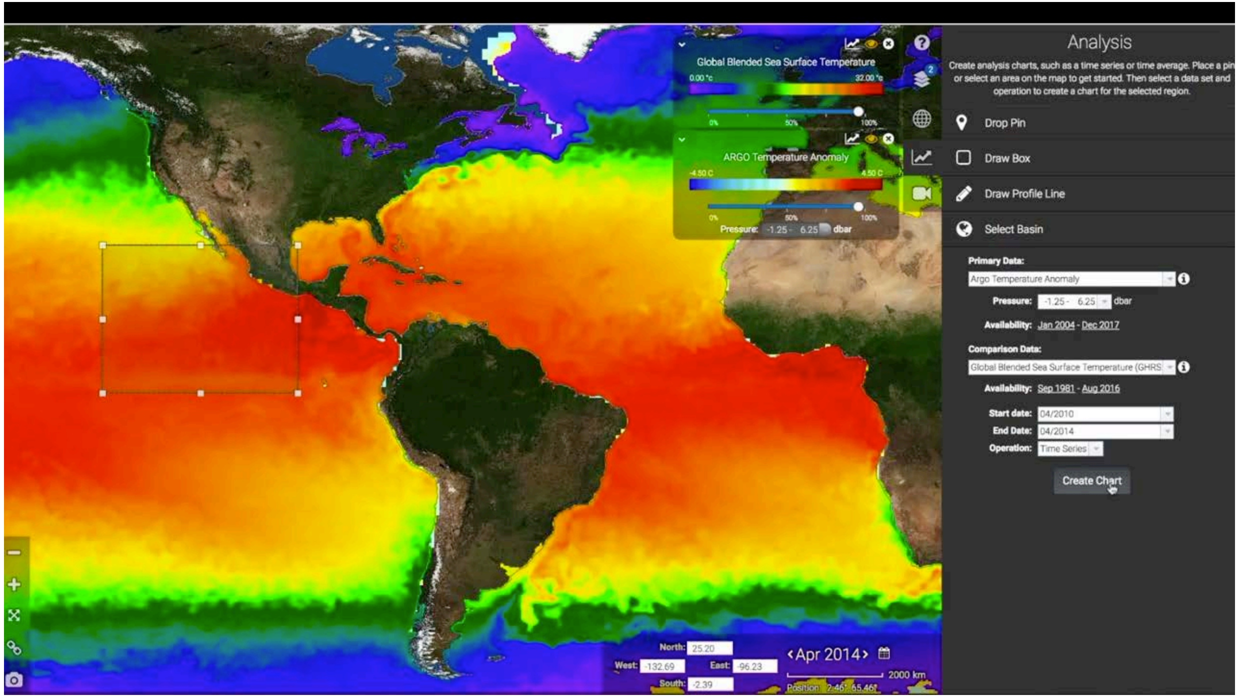
APACHE INCUBATOR

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Visualize and Analyze Sea Level

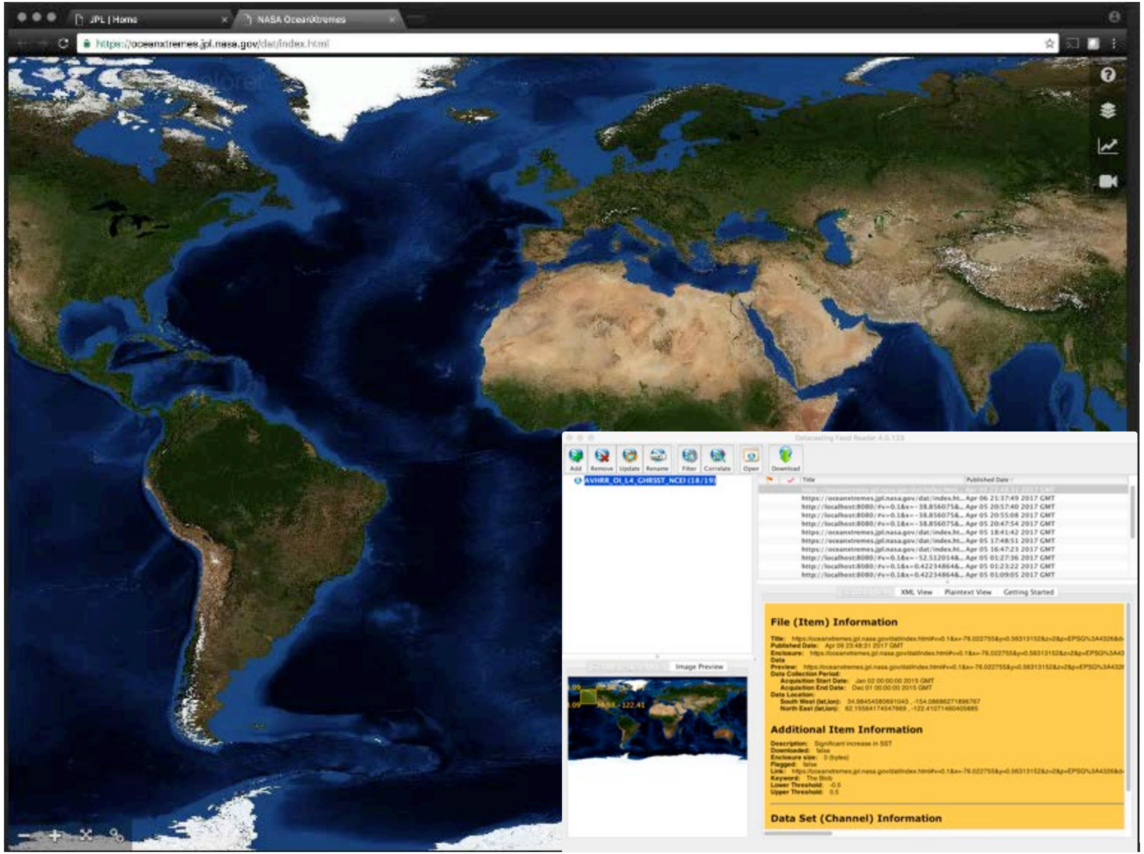
NASA Sea Level Change Science Team – Carmen Boening/PI - https://sealevel.nasa.gov



Analyze *in situ* and satellite observations

Analyze Sea Level
on mobiles

Analyze Ocean Anomaly – “The Blob”



- **Visualize** parameter
- **Compute** daily differences against climatology
- **Analyze** time series area averaged differences
- **Replay** the anomaly and visualize with other measurements
- **Document** the anomaly
- **Publish** the anomaly

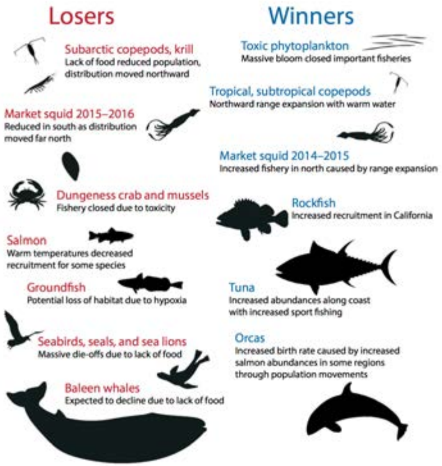
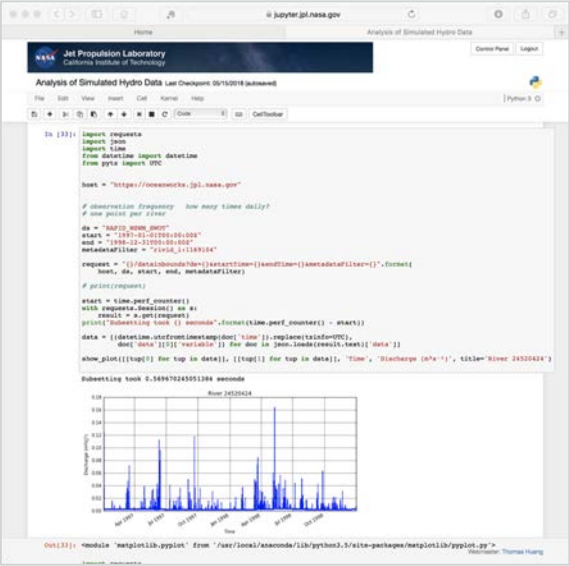
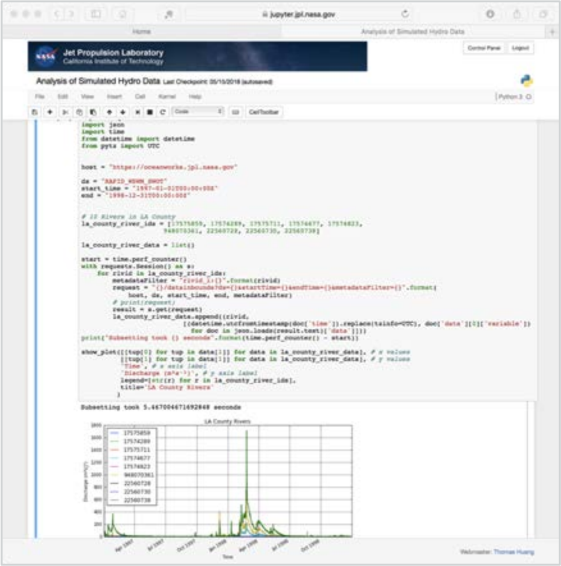


Figure from Cavole, L. M., et al. (2016). "Biological Impacts of the 2013–2015 Warm-Water Anomaly in the Northeast Pacific: Winners, Losers, and the Future." Oceanography 29.

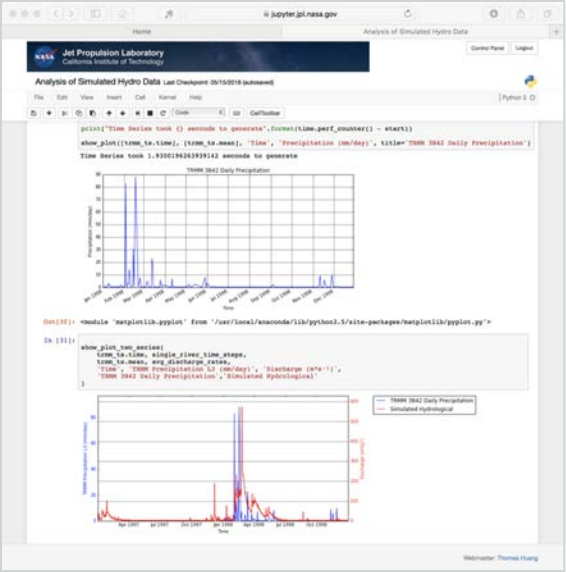
Hydrology



Retrieval of a single river time series



Retrieval of time series from 9 rivers



Time series coordination between TRMM and river

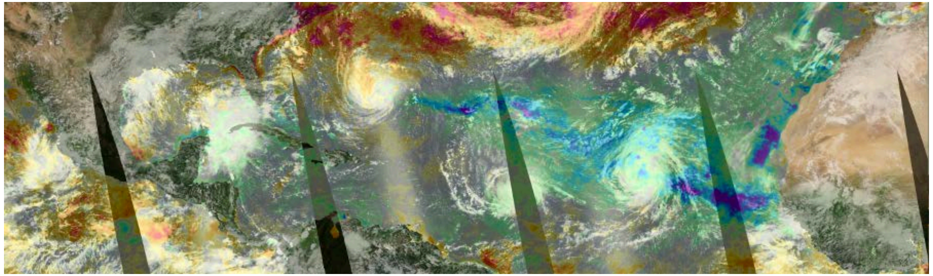
- Simulated hydrology data in preparation for SWOT hydrology
- **River data**: ~3.6 billion data points. 3-hour sample rate. Consists of measurements from ~600,000 rivers
- **TRMM data**: 17 years, .25deg, 1.5 billion data points
- Sub-second retrieval of river measurements
- On-the-fly computation of time series and generate coordination plot

In Situ to Satellite Matchup

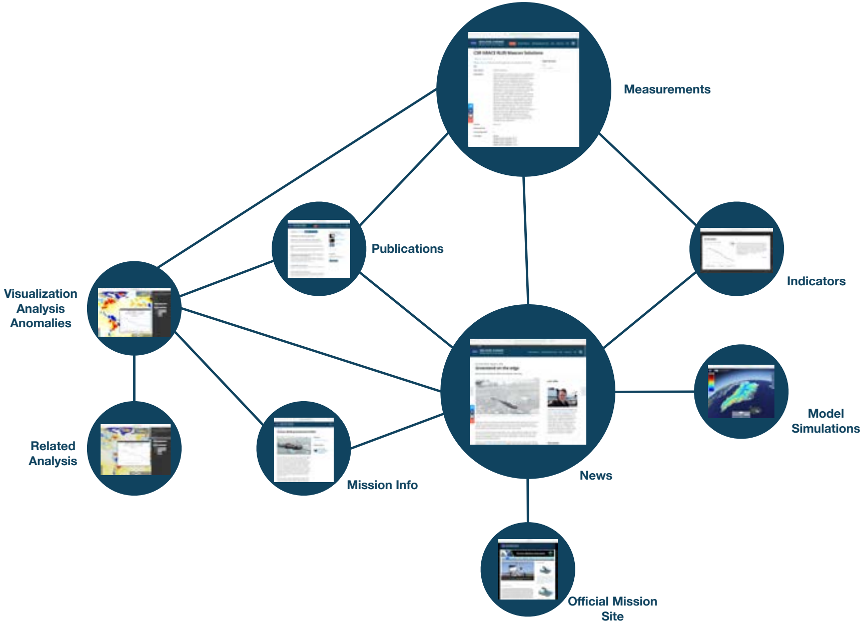
- Typically data matching is done using one-off programs developed at multiple institutions
- A primary advantage of SDAP's distributed matchup architecture is the reduction in duplicate development and man hours required to match satellite/in situ data
  - Removes the need for satellite and in situ data to be collocated on a single server
  - Systematically recreate matchups if either in situ or satellite products are re-processed (new versions), i.e., matchup archives are always up-to-date.
- In situ data nodes at JPL, NCAR, and FSU operational.
- Provides data querying, subset creation, match-up services, and file delivery operational.
- Plugin architecture for in situ data source using SDAP/EDGE, a open source implementation of Open Search

Tackling Information Discovery

- One of the big changes in Earth science is finding the relevant data and related online information
- We are developing smarter data search and discovery solution that is capable of adjusting search result according how user search, retrieval, and external events
- Use Machine Learning methods to adjust search ranking by taking a number of features into consideration
- Semantically mind dataset metadata to identify relationship
- Dynamically detect relationship between data, models, tools, publications, and news
- **Relevancy** is Domain-specific, Personal, Temporal, and Dynamic

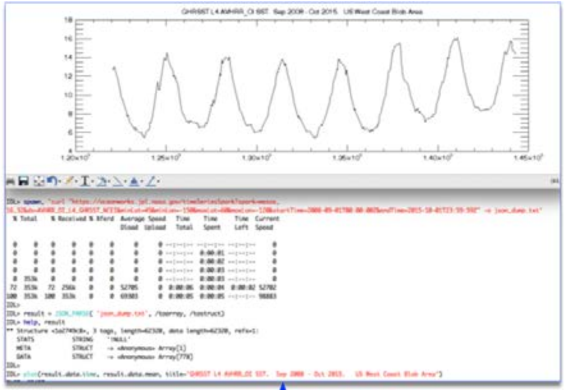


Air-sea Interaction during Hurricanes Florence, Joyce, and Helene in the Atlantic Ocean



Interact with Analytics Platform using any Programming Language

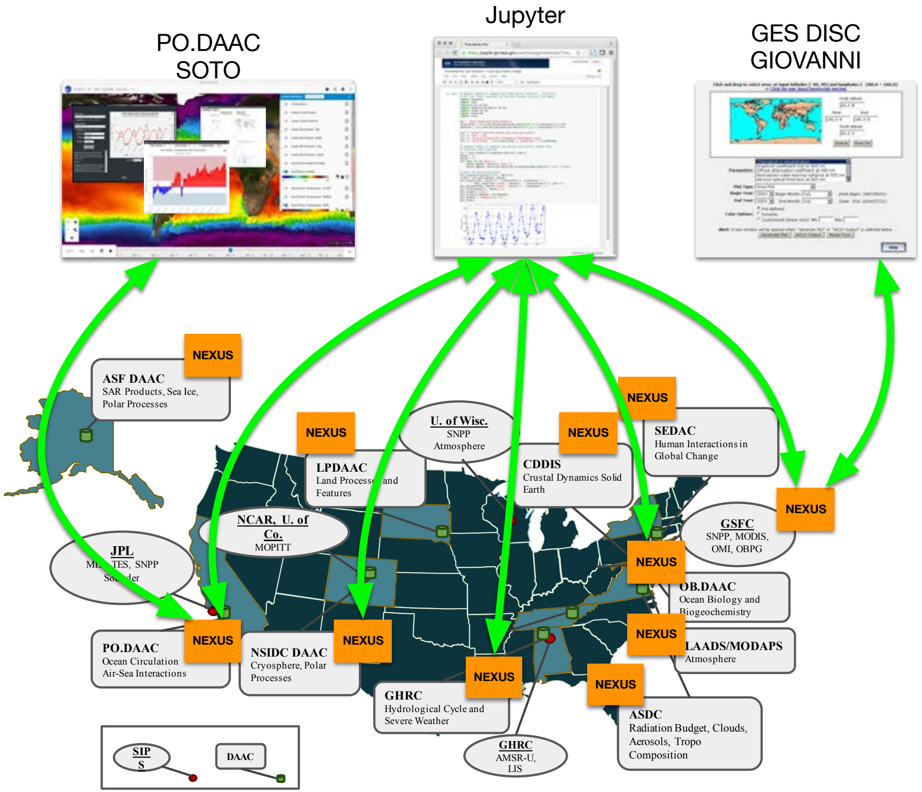
```
IDL> spawn, 'curl
"https://oceanworks.jpl.nasa.gov/timeSeriesSpark?spark=mesos,16,32&ds=AVHRR_OI_L4_
GHRSST_NCEI&minLat=45&minLon=-150&maxLat=60&maxLon=-120&startTime=2008-09-
01T00:00:00Z&endTime=2015-10-01T23:59:59Z" -o json_dump.txt'
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left  Speed

  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:--  0:00:01 --:--:--     0
  0     0    0     0    0     0      0      0 --:--:--  0:00:02 --:--:--     0
  0     0    0     0    0     0      0      0 --:--:--  0:00:03 --:--:--     0
  0  353k    0     0    0     0      0      0 --:--:--  0:00:03 --:--:--     0
 72  353k   72  256k    0     0  52705      0  0:00:06  0:00:04  0:00:02 52702
100  353k  100  353k    0     0  69303      0  0:00:05  0:00:05 --:--:-- 98883
IDL>
IDL> result = JSON_PARSE( 'json_dump.txt', /toarray, /tostruct)
IDL> help, result
** Structure <1a2749c8>, 3 tags, length=62320, data length=62320, refs=1:
   STATS            STRING    '!NULL'
   META             STRUCT    -> <Anonymous> Array[1]
   DATA             STRUCT    -> <Anonymous> Array[778]
IDL>
IDL> plot(result.data.time, result.data.mean, title='GHRSST L4 AVHRR_OI SST.  Sep
2008 - Oct 2015.   US West Coast Blob Area')
PLOT <29457>
```



Credit: Ed Armstrong
Jun. 05, 2018

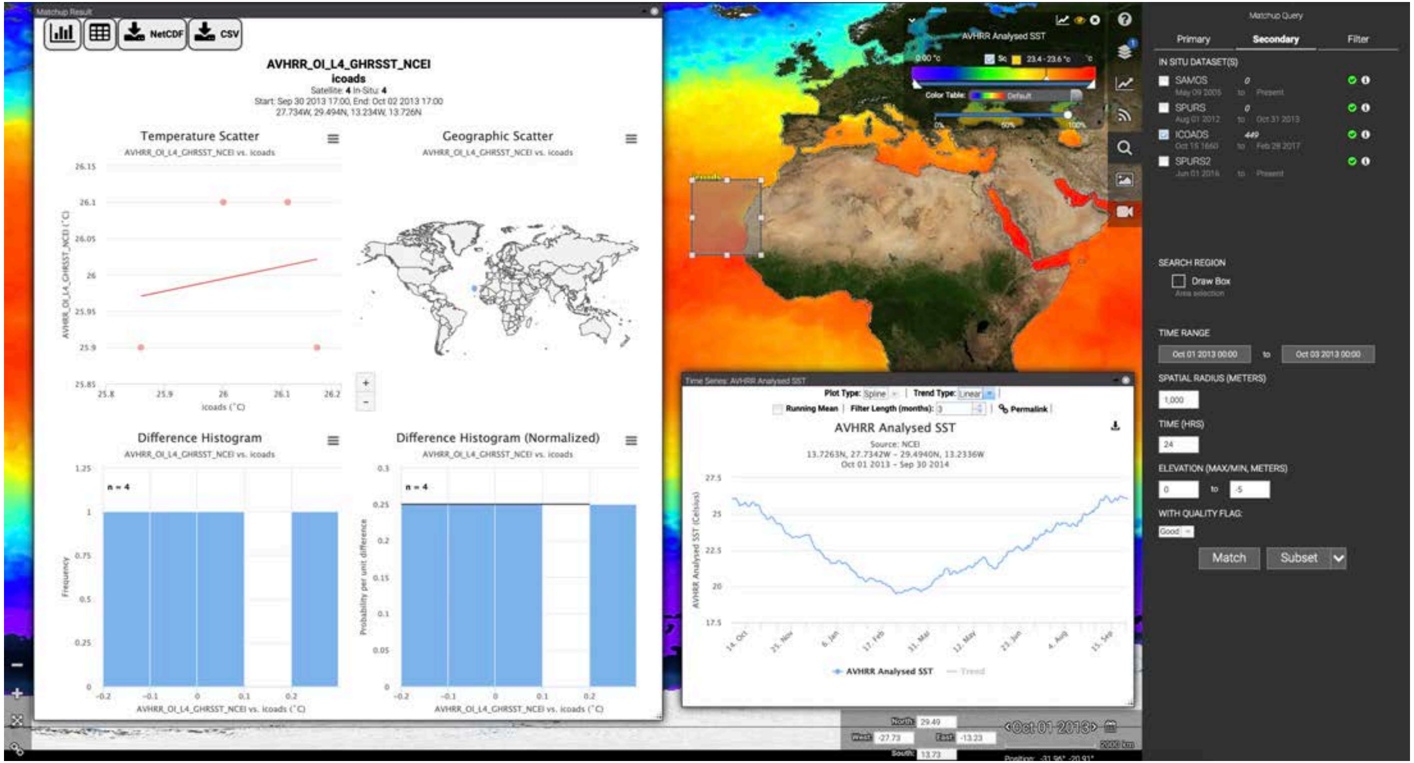
Multi-Variable Analysis

- Public accessible RESTful analytic APIs where computation is next to the data
- NEXUS as the analytic engine infused and managed by the DAACs on the Cloud
- Researchers can perform multi-variable analysis using any web-enabled devices without having to download files
- Reduce unnecessary egress charges
- An architecture to enable next generation of scientific applications



OceanWorks' Integrated OceanExplorer

Analytics and Distributed In situ to Satellite Matchup



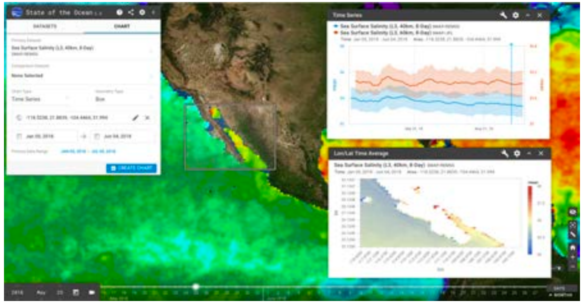
National Aeronautics and Space Administration

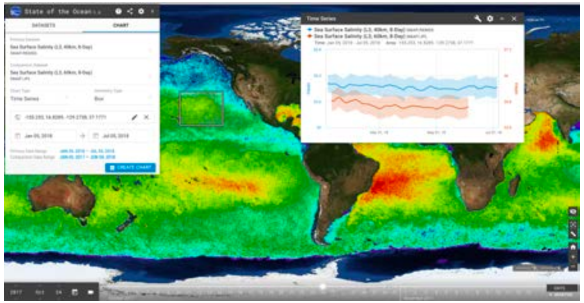
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

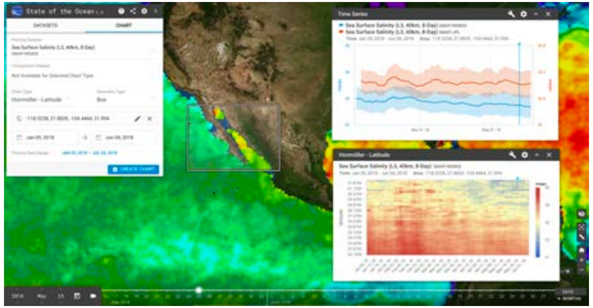
PO.DAAC's SOTO version 5.0

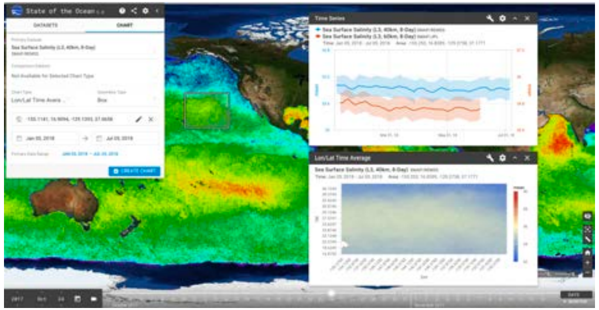
NASA Earth Science Data and Information System (ESDIS) – Suresh Vannan / PO.DAAC Manager

- NASA's Physical Oceanography Distributed Active Archive Center (PO.DAAC) is an element of the Earth Observing System Data and Information System (EOSDIS)
- PO.DAAC's mission is to preserve NASA's ocean and climate data and make these universally accessible and meaningful
- State of the Ocean (SOTO) is a PO.DAAC's popular visualization tool for the physical oceanography community
- SOTO v5 will be integrated with Apache SDAP and operate on the Amazon Cloud for on-the-fly data analytics
- Certified by PO.DAAC's System Integration Test (SIT). Testing under PO.DAAC's User Acceptance Testing (UAT) in preparation for its June 25, 2019 Annual User Working Group (UWG)









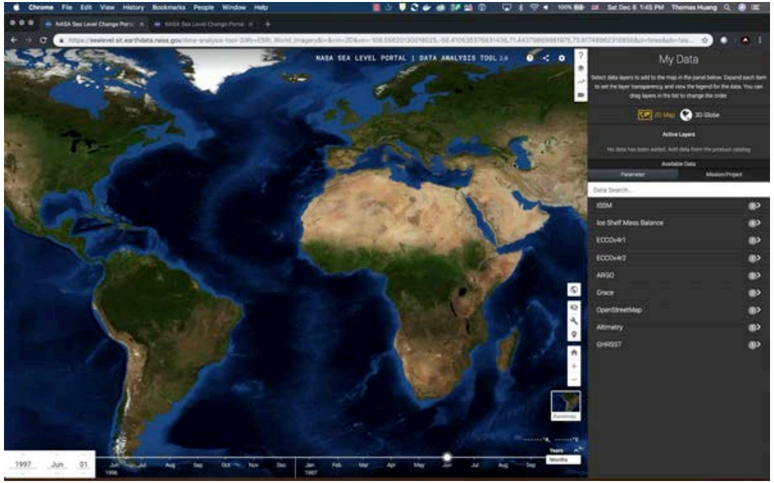
https://podaac.jpl.nasa.gov

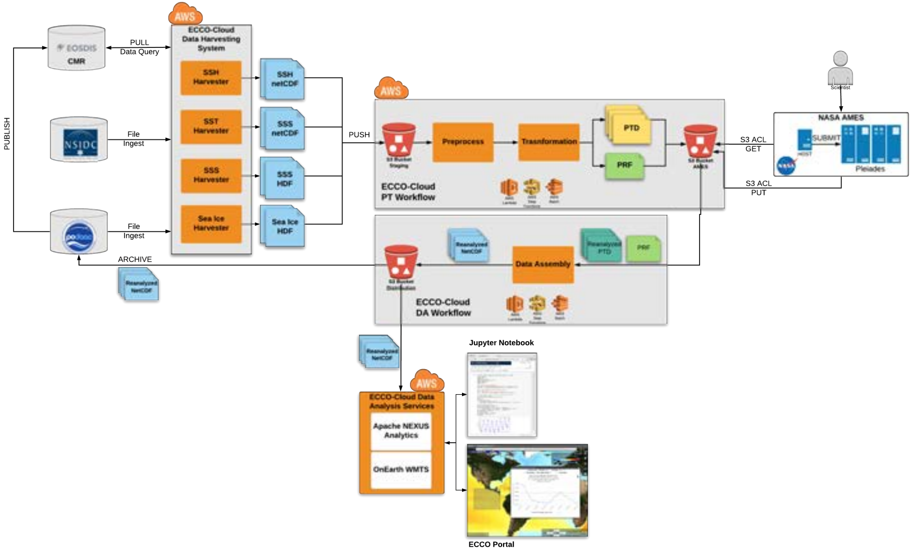
National Aeronautics and Space Administration

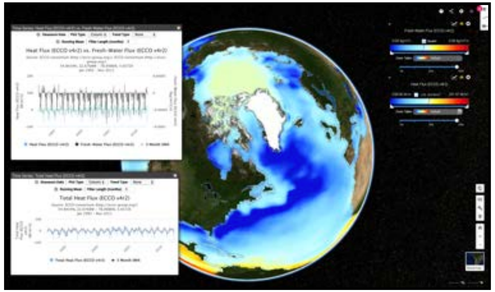
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Estimating the Circulation and Climate of the Ocean

NASA ACCESS Program – Patrick Heimbach/PI







- The **Estimating the Circulation and Climate of the Ocean (ECCO)** global ocean state estimation system (https://ecco.jpl.nasa.gov) is the premier tool for synthesizing NASA's diverse Earth system observations into a complete physical description of Earth's time-evolving full-depth ocean and sea ice system.
- Automate generation of ECCO reanalysis products into CF-compliant netCDF products
- Integrating Amazon Cloud with NASA Ames Pleiades petascale supercomputer
- Establish ECCO Data Analysis Services and web portal for interactive visualization and analysis, and distribution
- Support multi-dimensional data visualization and analysis

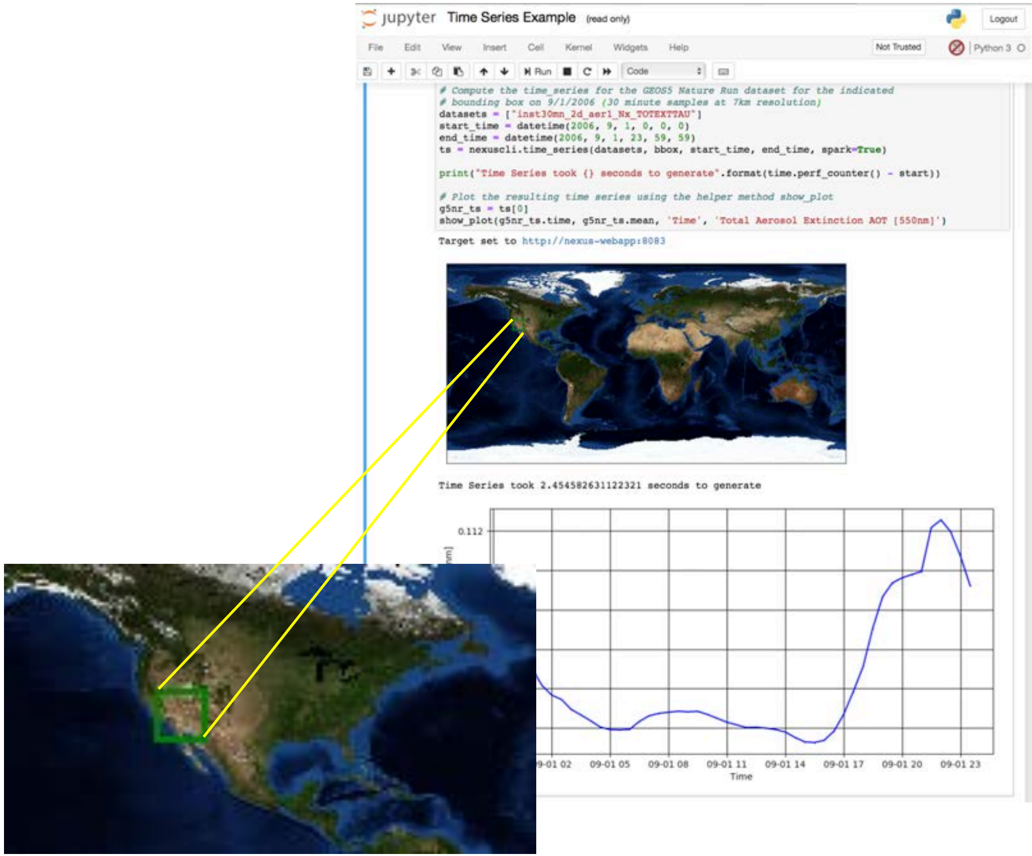
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Notebook Analytics: Parallel Climate Model Evaluation

NASA CMAC Program – Brian Wilson/PI

- To further our understanding of climate change by advanced analytics in a convenient notebook interface.
- Further improve SDAP/NEXUS performance of gridded products
- <u>Right panel</u>: Jupyter notebook prototype demonstration of rapid subsetting and time series calculation with SDAP/NEXUS.
- GEOS5 Nature Run (G5NR)
    - 7 km resolution
    - 30 minute model outputs
    - Plot of total aerosol extinction AOT [550 nm] with 30 minute samples
- Time series of 10 degree bounding box subset for 9/1/2006 took ~2 s on an M5.2xlarge AWS EC2 instance.
- Plan to scale for full 2-year simulation
    - June 2005 – June 2007
    - ~4 PB (44 variables)



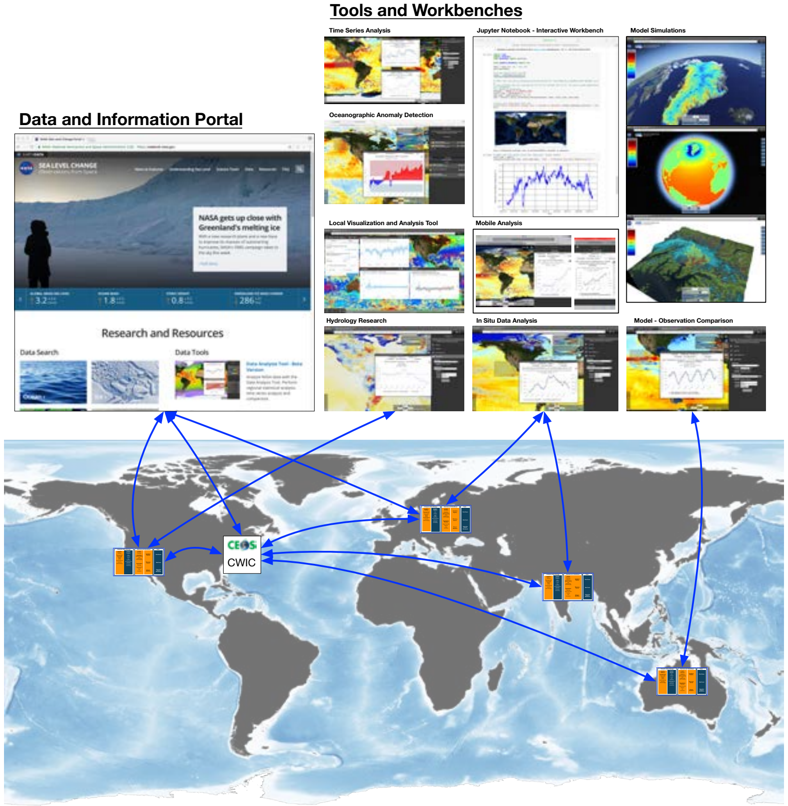
National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

COVERAGE's Distributed Analytics Center Architecture

NASA Physical Oceanography Program – Vardis Tsontos/PI

- **Committee of Earth Observation Satellites (CEOS) Ocean Variables Enabling Research and Applications for GEO (COVERAGE) Initiative**
- Seeks to provide **improved access** to **multi-agency ocean remote sensing data** that are **better integrated with in-situ and biological observations**, in support of **oceanographic and decision support applications** for societal benefit.
- A community-support open specification with common taxonomies, information model, and API (maybe security)
- Putting value-added services next to the data to eliminate unnecessary data movement
- Avoid data replication. Reduce unnecessary data movement and egress charges
- Public accessible RESTful analytic APIs where computation is next to the data
- Analytic engine infused and managed by the data centers perhaps on the Cloud
- Researchers can perform multi-variable analysis using any web-enabled devices without having to download files

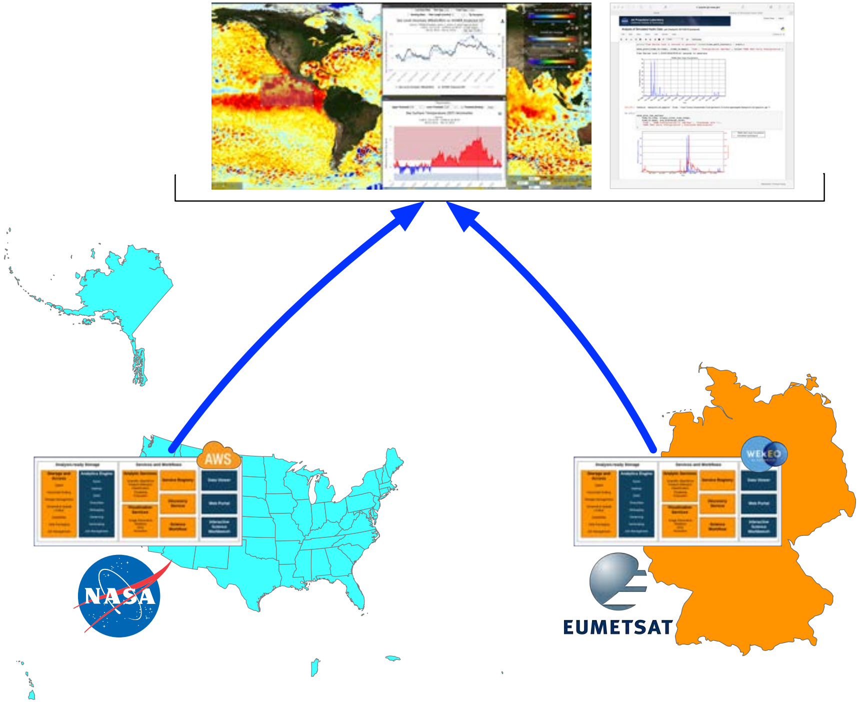


Credit: T. Huang, V. Tsontos, J. Vazquez, M. Chin - JPL

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

COVERAGE Phase B

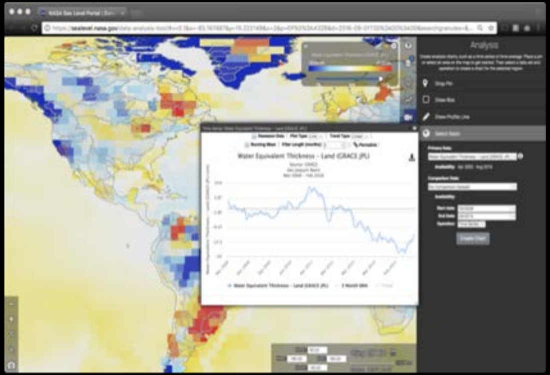


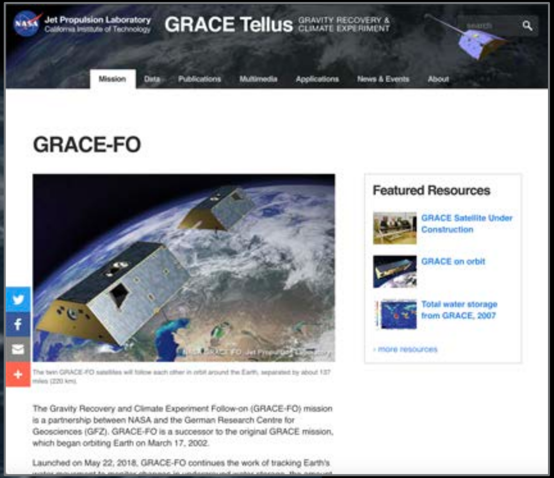
- WEkEO
  - Copernicus Data and Information Access Services (DIAS)
    1. Copernicus Data
    2. Virtual Environment and Tools
    3. User Support
  - Harmonized Data Access for Satellite data and Services
  - Virtualized infrastructure for personal sandboxes
  - Pre-configured tools
- COVERAGE Phase B
  - Establish US Node on Amazon Cloud
  - Establish EU Node on WEkEO at EUMETSAT
  - Establish COVERAGE data portal and analysis tool powered by the COVERAGE Nodes at US and EU

GRACE-FO Portal and Data Analysis Tool

GRACE/GRACE-FO Science - Felix Landerer/PI

- We are developing the new GRACE Follow-On Data Science Portal – https://grace.jpl.nasa.gov
- Goals
  - Common information model
  - Unified data search and access
  - Automated, serverless data processing, analysis and image generation system
  - Integrated with Google Analytics
  - New scientific data analytics capabilities
    - Hydrological basin analysis
    - Regional – country, continent, ocean basin, etc.
    - Multivariate data analysis
  - Deploy on Amazon Web Service with auto-scaling

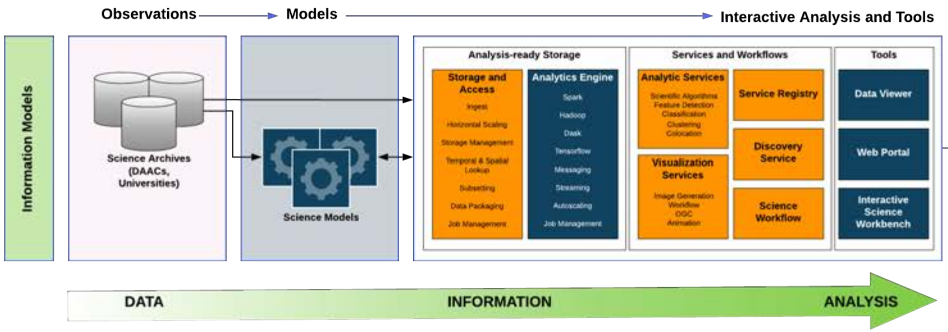


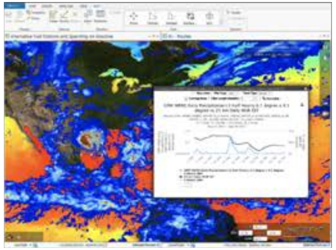


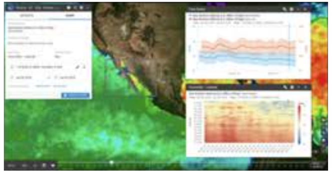
Integration with ESRI ArcGIS

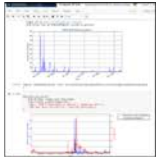
NASA AIST – Thomas Huang/PI

- Enable scientists to access and analyze NASA remote sensing data directly within the popular ArcGIS application
- Benefits
  - Integration to a popular commercial GIS desktop platform
  - Improve access of physical oceanography data
  - Enable large-scale, multivariate analysis without massive file downloads
  - Improve application of SWOT data for hydrological research
- The plugin developed from this effort will be open sourced as part of the Apache SDAP. It will eliminate the need for data provider to host an ESRI product
- Integration with OnEarth (WMTS) open source visualization service. The same service used by the NASA GIBS
- Promote use and access of NASA EO data by actively engaging the ESRI user community







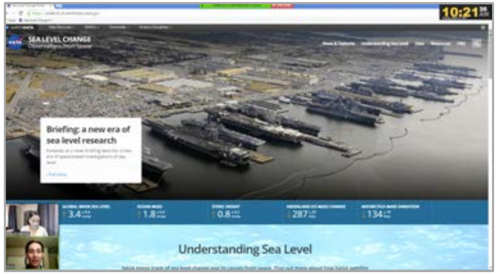


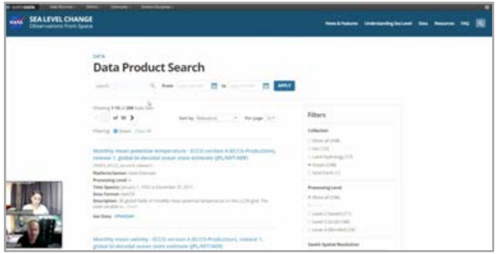
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Know The User's Real Needs

- **Work on improving communication - building bridge between IT and science**
  - **JPL's Data Science Program** is consists of technologists, project scientists, mission operations, etc.
  - Our science users tends get overwhelmed by tech jargons and cloud terminology
  - Learn to develop common language
- **Understand** how and for what purposes users obtain data and information
- **Describe** users' pain points and unmet needs for extracting, visualizing, comparing and analyzing science data
- **Identify** architectural approaches for tackling the real needs and identify opportunities for enhancing cross-disciplinary collaborative activities on the web portal.





Building Community-Driven Open Data and Open Source Solutions

- Deliver solutions to establish coherent platform solutions
- Embrace open source software
- Community validation
- Evolve the technology through community contributions
- Share recipes and lessons learned
- Technology demonstrations
- Host webinars, hands-on cloud analytics workshops and hackathons

2019 EGU – NASA Hyperwall

Big Data Analytics and Cloud Computing Workshop, 2017 ESIP Summer Meeting, Bloomington, IN

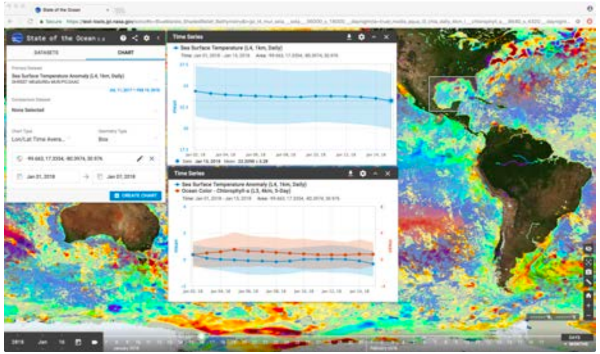
2019 JPL Data Science Showcase

Partner with NASA and non-NASA Projects - Deliver to Production

- The gap between visionary to pragmatists is significant. It must be the primary focus of any long-term high-tech marketing plan – Geoffrey Moore
- Become an expert in the production environment and devote resources in creating automations
- Give project engineering team early access to the PaaS
- Deliver all technical documents and work with project system engineering
- Provide user-focused trainings

NASA's Sea Level Change Team

CEOS SIT Technical Workshop



NASA's Physical Oceanography Distributed Active Archive Center (PO.DAAC)

Presentation | Evaluation | Collaboration

Presentation | Evaluation | Collaboration

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

In Summary

- You've got to think about big things while you're doing small things, so that all the small things go in the right direction – Alvin Toffler
- Focus on end-to-end data and computation architecture, and the total cost of ownership
- JPL Strategy is to drive Data Science into the fabric of JPL by
    - Launching cross-institution pilots
    - Building a trained workforce
    - Linking to the mission-science data lifecycle
- Invest in Interactive Analytics that simplifies the integration of *multiple* Earth observing remote sensing instruments; comparison against models
- AI and Data Science will be an essential part of NASA's future!

Thomas Huang

thomas.huang@jpl.nasa.gov

Jet Propulsion Laboratory

California Institute of Technology

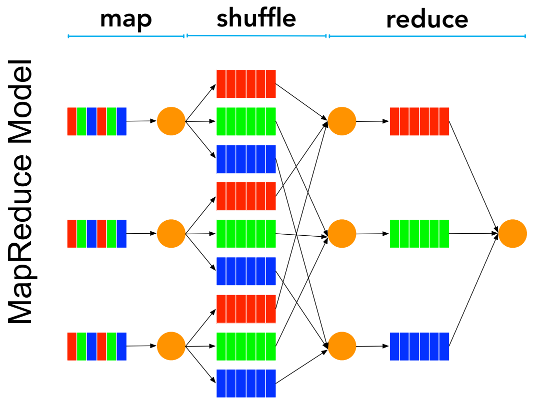
BACKUP

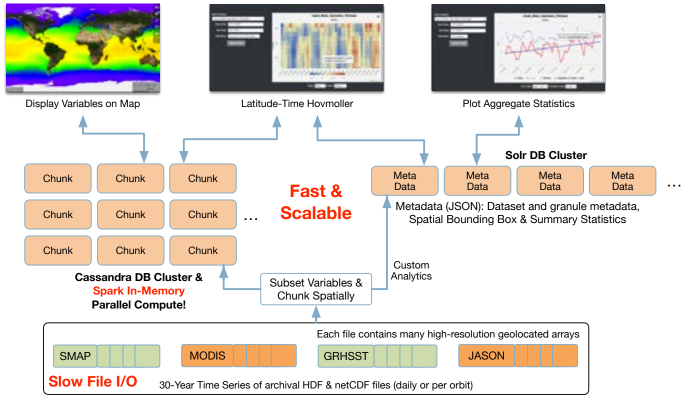
National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Scalable Data Analytic Solution

- **MapReduce**: A programming model for expressing distributed computations on massive amount of data and an execution framework for large-scale data processing on clusters of commodity servers. - J. Lin and C. Dyer, "*Data-Intensive Text Processing with MapReduce*"
    - **Map**: splits processing across cluster of machines in parallel, each is responsible for a record of data
    - **Reduce**: combines the results from Map processes
- **SDAP's NEXUS** is a data-intensive analysis solution using a new approach for handling science data to enable large-scale data analysis
    - Streaming architecture for horizontal scale data ingestion
    - Scales horizontally to handle massive amount of data in parallel
    - Provides high-performance geospatial and indexed search solution
    - Provides tiled data storage architecture to eliminate file I/O overhead
    - A growing collection of science analysis webservices

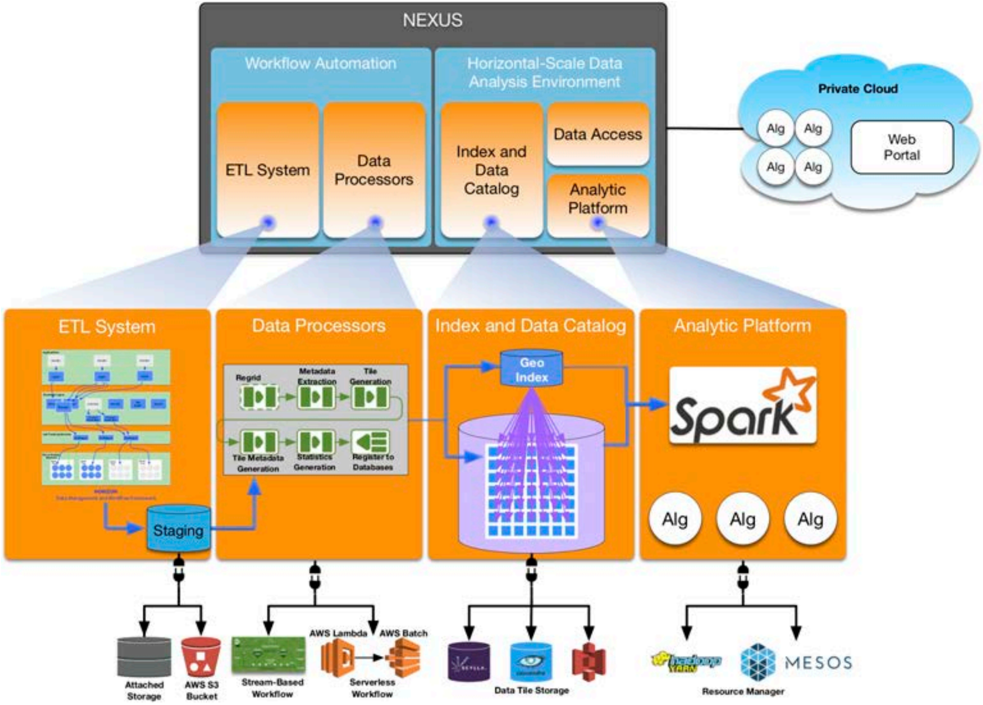




Two-Database Architecture

Evolve the Parallel Analytics Architecture

- **Several container-based deployment options**
  - Local on-premise cluster
  - Private Cloud (OpenStack)
  - Amazon Web Service
- **Automate Data Ingestion with Image Generation**
  - Cluster based
  - Serverless (Amazon Lambda and Batch)
- **Data Store Options**
  - Apache Cassandra
  - ScyllaDB
  - Amazon Simple Storage Service (S3)
- **Resource Management Options**
  - Apache YARN
  - Apache MESOS
- **Analytic Engine Options**
  - Custom Apache Spark Cluster
  - Amazon Elastic MapReduce (EMR)
  - Amazon Athena (work-in-progress)



Apache SDAP's NEXUS supports public/private Cloud and local cluster deployments

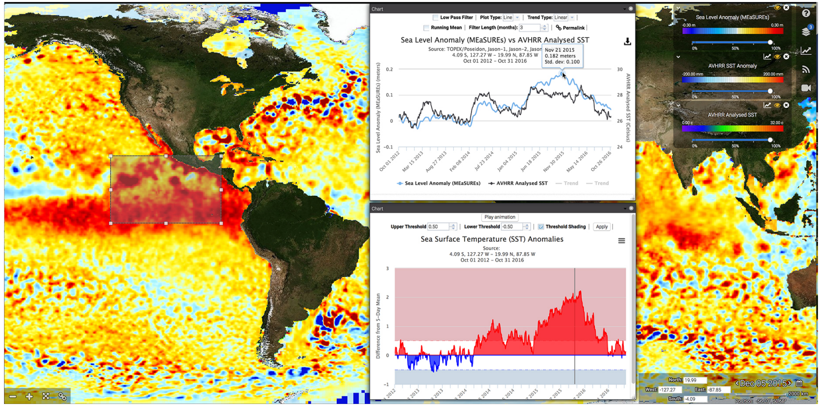
Opportunities Enabled by Data Science

1. Support scalability to capture and analyze NASA observational data

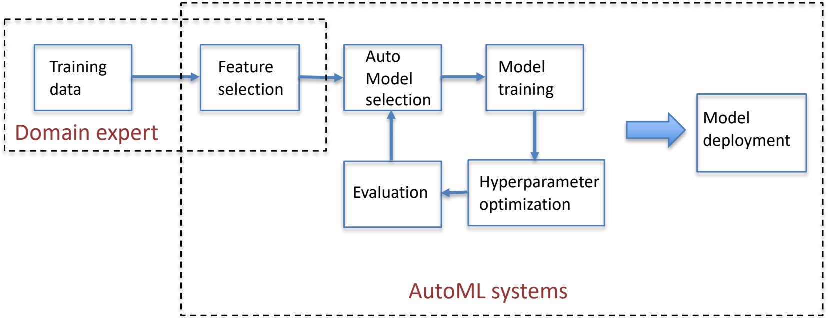
2. Apply data-driven approaches across the entire data lifecycle

3. Increase access, integration and use of highly distributed archival data

4. Increased data science services for on-demand, interactive visualization and analytics



NASA AIST: OceanXtremes - Anomaly Detection Solution



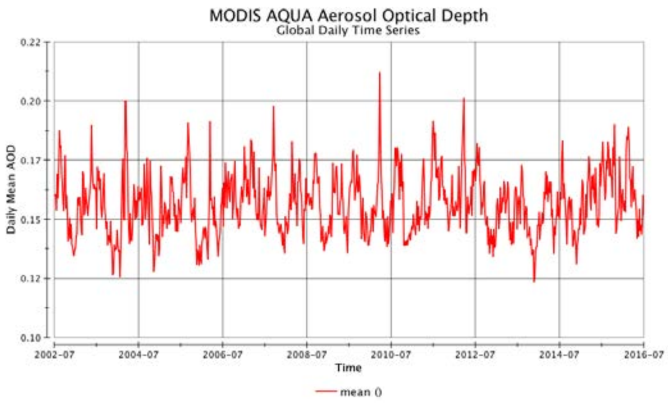
Automate Machine Learning

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Parallel Analytics Performance

Dataset: MODIS AQUA Daily
Name: Aerosol Optical Depth 550 nm (Dark Target) (MYD08_D3v6)
File Count: 5106
Volume: 2.6GB
Time Coverage: July 4, 2002 – July 3, 2016

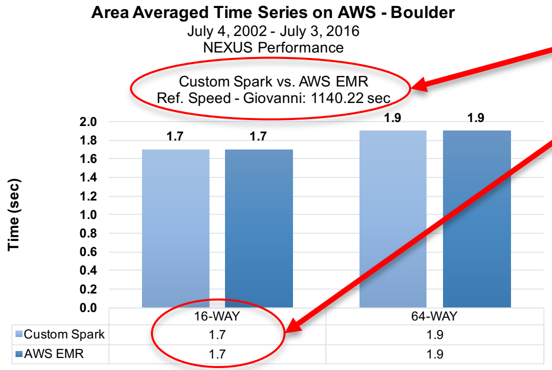


Giovanni: A web-based application for visualize, analyze, and access vast amounts of Earth science remote sensing data without having to download the data.

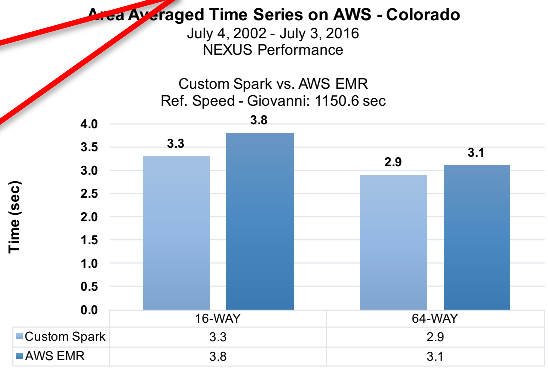
- Represents current state of data analysis technology, by processing one file at a time
- Backed by the popular NCO library. Highly optimized C/C++ library

AWS EMR: Amazon’s provisioned MapReduce cluster

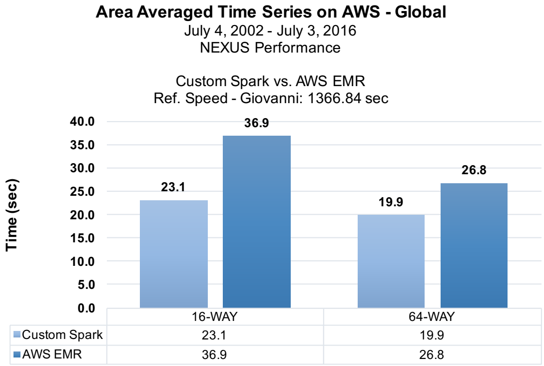
Giovanni: 20 min
NEXUS: 1.7 sec



	16-WAY	64-WAY
Custom Spark	1.7	1.9
AWS EMR	1.7	1.9



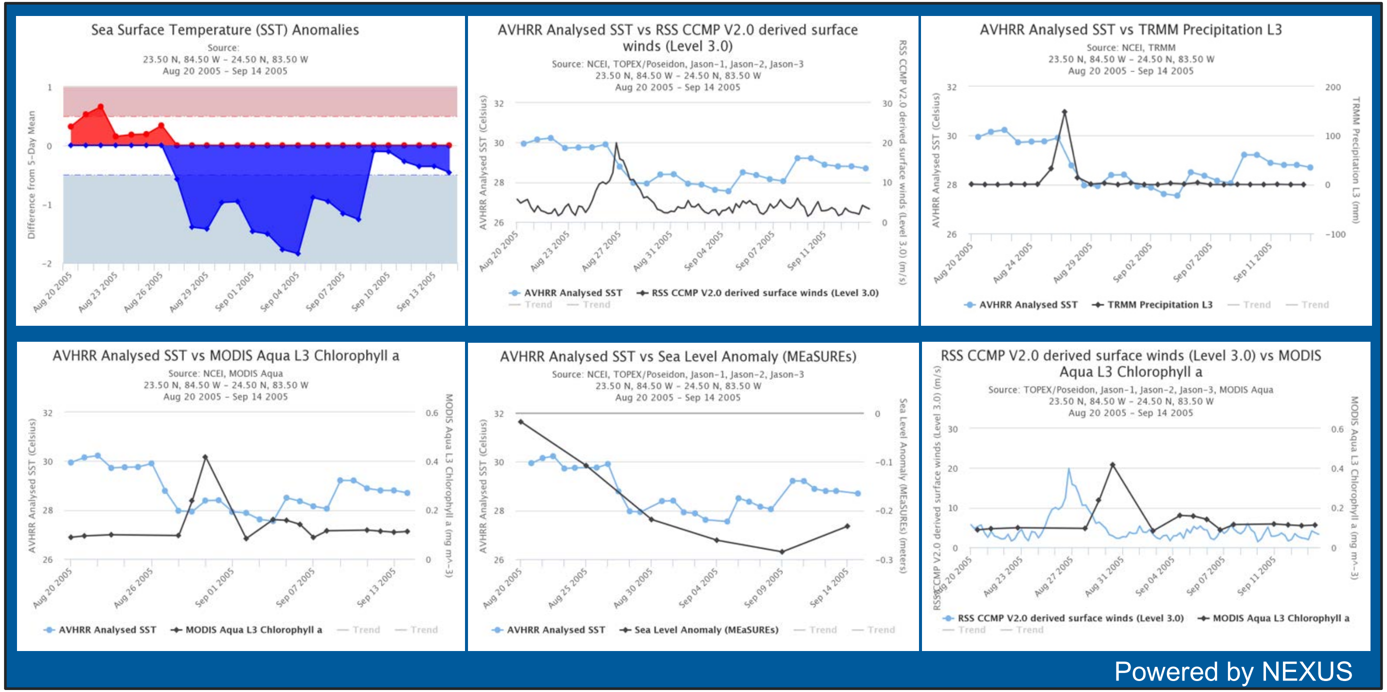
	16-WAY	64-WAY
Custom Spark	3.3	2.9
AWS EMR	3.8	3.1



	16-WAY	64-WAY
Custom Spark	23.1	19.9
AWS EMR	36.9	26.8

Algorithm execution time. Excludes Giovanni’s data scrubbing processing time

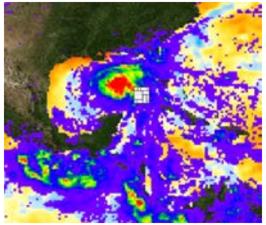
Hurricane Katrina Study



A study of a Hurricane Katrina–induced phytoplankton bloom using satellite observations and model simulations
Xiaoming Liu, Menghua Wang, and Wei Shi
JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, C03023, doi:10.1029/2008JC004934, 2009

Hurricane Katrina passed to the southwest of Florida on Aug 27, 2005. The ocean response in a 1 x 1 deg region is captured by a number of satellites. The initial ocean response was an immediate cooling of the surface waters by 2 ℃ that lingers for several days. Following this was a short intense ocean chlorophyll bloom a few days later. The ocean may have been "preconditioned' by a cool core eddy and low sea surface height.

The SST drop is correlated to both wind and precipitation data. The Chl-A data is lagged by about 3 days to the other observations like SST, wind and precipitation.



Hurricane Katrina TRMM overlay SST Anomaly